



Figure 1. N-linked oligosaccharides for glycosylation

Peptide sequence (SEQ ID NO:5):

MPFVNKQFNY KDPVNGVDIA YIKIPNAGQM QPVKAFKIHN KIWVIPERDT FTNP EEGDLN
 PPPEAKQVPV SYDSTYLST DNEKDNLYKG VTKLFERIYS TDLGRMLLTS IVRGIPFWGG
 STIDTELKVI DTNCINVIQP DGSYRSEELN LVIIGPSADI IQFECKSFGH EVLNLTRNGY
 GSTQYIRFSP DFTFGFEESL EVDTNPLLGA GKFATDPAVT LAHELHAGH RLYGIAINPN
 RVFKVNTNAY YEMSGLEVSE EELRTFGGHD AKFIDSLQEN EFRLYYYNKF KDIASLTNKA
 KSIVGTTASL QYMKNVFKEK YLLSEDTSGK FSVDKLKFDK LYKMLTEIYT EDNFVKFFKV
 LNRKTYLNFD KAVFKINIVP KVNYYTYDGF NLRNTNLAAN FNGQNTTEINN MNFTKLKNFT
 GLFEFYKLLC VRGIITSKTK SLDKGYNKAL NDLCKVNNW DLFFSPSEDN FTNDLNKGEE
 ITSNTNIEAA EENISLDLIQ QYYLTNFNFDN EPENISLENL SSDIIGQLEL MPNIERFPNG
 KKYELDQYTM FHYLRAQEFH HGKSRIALTN SVNEALLNPS RYVYFFSSDY VKKVNKATEA
 AMFLGWVEQL VYDFTDETSE VSTTDKIADI TIIIPYIGPA LNIGNMPLYKD DFGALIFSG
 AVILLEFIPE IAIPVLGTFA LVSYIANKVL TVQTIDNALS KRNEKWDEVY KYIVTNWLAK
 VNTQIDLIRK KMKEALENQA EATKAIINYQ YNQYTEEEKN NINFNIDDL SKLNESINKA
 MININKFLNQ CSVSYLMNSM IPYGVKRLD FDSLKDALL KYIYDNRGTL IGQVDRKDK
 VNNTLSTDIP FQLSKYVDNQ RLLSTFTEYI KNIINTSILN LRYESNHLID LSRYSKINI
 GSKVNFDPID KNQIQLFNLE SSKIEVILKN AIVYNSMYEN FSTFSWIRIP KYFNSISLNN
 EYTIINCMEN NSGWKVS LNY GEIWTLODT QEIKQRVVFK YSQMINISDY INRWIFVTIT
 NNRLNNSKIY INGRLLDQKP ISNLGNIHAS NNIMFKLDGC RDTHRYIWIY YFNLFDKELN
 EKEIKDLYDN QSNNGILKDF WGDYLYQDKP YMLNLYDPN KYVDVNNVGI RGYMYLKGPR
 GSVMTTNIYL NSSLYRGTFE IIKKYASGNK DNIVRNNDRV YINVVVKNKE YRLATNASQA
 GVEKILSALE IPDVGNLSQV VVMKSKNDQG ITNKCKMNLQ DNNGNDIGFI GFHQFNNAIK
 LVASNWNRYQ IERSSRTLGC SWEFIPVDDG WGERPL

Peptides containing the motif 'N-X-S/T/C (X not P)'(underlined):

position	Peptide	SEQ ID NO:
167-177	SFGHEVL <u>NL</u> TR	6
382-393	V <u>NY</u> TYDGFNLR	7
394-415	NTNLAANFNGQNTTEINN <u>MN</u> FTK	8
418-427	<u>NFT</u> GLFEFYK	9
457-477	VNNWDLFFSPSEDN <u>FT</u> NDLN K	10
478-536	GEEITSNTNIEAAEENISLD LIQQYYLTNFNFDNEP <u>EN</u> ISI	11
773-779	<u>LN</u> ESINK	12
787-806	FLNQCSVSYLMNSMIPYGVK	13
841-855	VNNTLSTDIPFQLSK	14
872-882	NI <u>INT</u> SILNLR	15
930-948	NAIVYNSMYEN <u>F</u> STFSWIR	16
952-975	YFNSISLNNEYTIINCMENN <u>S</u> GWK	17
1001-1013	YSQMIN <u>IS</u> DYINR	18
1024-1028	<u>LN</u> NSK	19
1086-1098	DLYDNQSNNGILK	20
1141-1156	GSVMTTNIYL <u>N</u> SSLYR	21
1193-1204	LATNASQAGVEK	22
1205-1224	ILSALEIPDVGN <u>L</u> SQVVVMK	23

Figure 1 Continue.

Peptide sequence (SEQ ID NO:39):

KTKSLDKGYN KALNDLCIKV NNWDLFFSPS EDNFTNDLNK GEEITSDTNI EAAEENISLD
LIQQYYLTFN FDNEPENISI ENLSSDIIGQ LELMPNIERF PNGKKYELDK YTMFHYLRAQ
EFEHGKSRIA LTNSVNEALL NPSRVYTFFS SDYVKKVNKA TEAAMFLGWV EQLVYDFTDE
TSEVSTTDKI ADITIIIPYI GPALNIGNML YKDDFVGALI FSGAVILLEF IPEIAIPVLG
TFALVSYIAN KVLTVQTIDN ALSKRNEKWD EVYKYIVTNW LAKVNTQIDL IRKKMKEALE
NQAEATKAI NYQYNQYTEE EKNNINFNID DLSSKLNESI NKAMININKF LNQC SVSYLM
NSMIPYGVKR LEDFDASLKD ALLKYIYDNR GTLIGQVDRL KDKVNNTLST DIPFQLSKYV
DNQRLSTFT EYIKNIINTS ILNLRYESNH LIDLSRYASK INIGSKVNFD PIDKNQIQLF
NLESSKIEVI LKNAIVYNSM YENFSTSFWI RIPKYFNSIS LNNEYTIINC MENNSGWKVS
LNYGEIIWTL QDTQEIKQRV VFKYSQMINI SDYINRWIFV TITNNRLNNS KIYINGRLID
QKPISNLGNI HASNNIMFKL DGCRDTHRYI WIKYFNLFDK ELNEKEIKDL YDNQSN SGIL
KDFWGDYLQY DKPYMYMLNLY DPNKYVDVNN VGIRGYMYLK GPRGSMVTN IYLNSSLYRG
TKFIIKKYAS GNKDNIVRNN DRVYINVVVK NKEYRLATNA SQAGVEKILS ALEIPDVGNL
SQVVVMKSKN DQGITNKCKM NLQDNNGNDI GFIGFHQFNN IAKLVASNWY NRQIERSSRT
LGCSWEFIPV DDGWERPL

Peptides containing the motif 'N-X-S/T/C (X not P)':

position	peptide	SEQ ID NO:
20-40	VNNWDLFFSPSEDNFTNDLN K	25
41-99	GEEITSDTNIEAAEENISLD LIQQYYLTFNFDNEPENISI ENLSSDIIGQLELMPNIER	26
336-342	LNESINK	27
350-369	FLNQCSVSYLMNSMIPYGVK	28
404-418	VNNTLSTDIPFQLSK	29
435-445	NIINTSILNLR	30
493-511	NAIVYNSMYENFSTSFWIR	31
515-538	YFNSISLNNEYTIINCMENN SGWK	32
564-576	YSQMINISDYINR	33
587-591	LNNSK	34
649-661	DLYDNQSN SGILK	35
704-719	GSVMTTNIYLNSSLYR	36
756-767	LATNASQAGVEK	37
768-787	ILSALEIPDVGNLSQVVVMK	38

Figure 2.

Peptide sequence (SEQ ID NO: 39):

MPFVNKQFNY KDPVNGVDIA YIKIPNAGQM QPVKAFKIHN KIWVIPERDT FTNPEEGDLN
 PPPEAKQVPV SYDSTYLST DNEKDNYLKG VTKLFERIYS TDLGRMLLTS IVRGIPFWGG
 STIDTELKVI DTNCINVIQP DGSYRSEELN LVIIGPSADI IQFECKSFGH EVLNLTRNGY
 GSTQYIRFSP DFTFGFEESL EVDTNPLLGA GKFATDPAVT LAHELIHAGH RLYGIAINPN
 RVFKVNTNAY YEMSGLEVSF EELRTFGGHD AKFIDSLQEN EFRLYYYNKF KDIASLTNKA
 KSIVGTTASL QYMKNVFKEK YLLSEDTSKG FSVDKLFKFDK LYKMLTEIYT EDNFVKFFKV
 LNRKTYLNFD KAVFKINIVP KVNYTIYDGF NLRNTNLAAN FNGQNTTEINN MNFTKLKNFT
 GLFEFYKLLC VRGIITSKTK SLDKGYNKAL NDLCIKVNNW DLFFSPSEDN FTNDLNKGEE
 ITSNTNIEAA EENISLDLIQ QYYLTFNFDN EPENISIEENL SSDIIGQLEL MPNIERFPNG
 KKYELDKYTM FHYLRAQEFH HGKSRIALTN SVNEALLNPS RVYTFSSDY VKKVNKATEA
 AMFLGWVEQL VYDFTDETSE VSTTDKIADI TIIIPYIGPA LNIGNMLYKD DFGALIFSG
 AVILLEFIPE IAIPVLGTFA LVSYIANVKL TVQTIDNALS KRNEKWDEVY KYIVTNWLAK
 VNTQIDLRK KMKEALENQA EATKAIINYQ YNQYTEEEKN NINFNIDDL SKLNESINKA
 MININKFLNQ CSVSYLMNSM IPYGVKRLD FDASLKDALL KYIYDNRGTL IGQVDRKDK
 VNNTLSTDIP FQLSKYVDNQ RLLSTFTEYI KNIINTSILN LRYESNHLID LSRYASKINI
 GSKVNFDPID KNQIQLFNLE SSKIEVILKN AIVYNSMYEN FSTSFWRIP KYFNSISLNN
 EYTIINCMEN NSGWKVSILN GEIITWLQDT QEIKQRVVFK YSQMINISDY INRWIFVTIT
 NNRLNNSKIY INGRLLDQKP ISNLGNHIS NNIMFKLDGC RDTHRYIWIY YFNLFDEKLN
 EKEIKDLYDN QSNISGILKDF WGDYLYQYDKP YYMLNLYDPN KYVDVNNVGI RGYMYLKGPR
 GSVMTTNIYL NSSLYRGTKF IIKKYASGNK DNIVRNNDRV YINVVVKNE YRLATNASQA
 GVEKILSALE IPDVGNSQV VVMKSKNDQG ITNKCKMNLQ DNNGNDIGFI GFHQFNNAK
 LVASNWNRY IERSSRTLGC SWEFIPVDDG WGERPL

Peptides containing S or T (underlined):

position	peptide	SEQ ID NO:
49-66	DTFTNPEEGDLNPPPEAK	40
67-84	QVPVSYYDSTYLSTDNEK	41
90-93	GVTK	42
98-105	IYSTDLGR	43
106-113	MLLTSIVR	44
114-128	GIPFWGGSTIDTELK	45
129-145	VIDTNCINVIQPDGSR	46
146-166	SEELNLVIIGPSADIIQFEC K	47
167-177	SFGHEVLNLTR	48
178-187	NGYGSTQYIR	49
188-212	FSPDFTFGFEESLEVDTNPL LGAGK	50
213-231	FATDPAVTLAHELIHAGHR	51
245-264	VNTNAYYEMSGLEVSFEELR	52
265-272	TFGGHDAK	53
273-283	FIDSLQENEFR	54
292-299	DIASLTNK	55
302-314	SIVGTTASLQYMK	56
321-330	YLLSEDTSKG	57
331-335	FSVDK	58
344-356	MLTEIYTEDNFVK	59
365-371	TYLNFDK	60
382-393	VNYTIYDGFNL	61
394-415	NTNLAANFNGQNTTEINNMNF TK	62
418-427	NFTGLFEFYK	63
433-438	GIITSK	64
439-440	TK	
441-444	SLDK	66

Figure 2

Continue

457-477	VNNWDLFF <u>S</u> P <u>S</u> EDNFTNDLN K	67
478-536	GEEITSDT <u>N</u> IEAAEENISLD LIQQYYLT <u>F</u> NFDNEPENIS <u>I</u>	68
548-555	ENLSSDIIGQLELMPNIER	69
564-565	YTMFH <u>Y</u> LR	
566-581	<u>S</u> R	71
582-592	IALTNSVNEALLNPS <u>R</u>	72
597-626	VYTF <u>F</u> SSDYVK	73
627-649	ATEAAMFLGWVEQLVYDFTD ETSEVSTTDK	74
650-688	IADITIIIPYIGPALNIGNM LYK	75
689-701	DDFVGALIFSGAVILLEFIP EIAIPVLGT <u>F</u> FALV <u>S</u> YIANK	76
712-720	VLTVQTIDNAL <u>S</u> K	77
721-729	YIVT <u>N</u> WLAK	78
734-744	VNTQIDLR	79
745-759	EALNQAEAT <u>K</u>	80
760-772	AIINYQYNQYTEEEK	81
773-779	NNINFNIDDL <u>S</u> SK	82
787-806	LNE <u>S</u> INK	83
808-816	FLNQCSVSYL <u>M</u> NSMIPYGVK	84
828-836	LEDFDAS <u>L</u> K	85
841-855	GTLIGQVDR	86
862-871	VNNTLSTDIPFQL <u>S</u> K	87
872-882	LLSTFTEYIK	88
883-893	NIINTSILNLR	89
894-897	YESNHLIDL <u>S</u> R	90
898-903	YAS <u>K</u>	91
912-923	INIG <u>S</u> K	92
930-948	NQIQLFNLE <u>S</u> SK	93
952-975	NAIVYNSMYENFSTSFWIR	94
976-994	YFNSISLNNEYTIINCMENN <u>S</u> GWK	95
1001-1013	VSLNYGEIIWTLQDTQEI <u>K</u>	96
1014-1023	YSQMINISDYINR	97
1024-1028	WIFVTITNNR	98
1035-1056	LNN <u>S</u> K	99
1062-1065	LIDQKPISNLGNIHASNNIM FK	100
1066-1098	DTHR	101
1141-1156	DLYDNQSN <u>S</u> GILK	102
1157-1159	GSVMTTNIYLN <u>S</u> SLYR	
1165-1170	GTK	104
1193-1204	YASGNK	103
1205-1224	LATNASQAGVEK	70
1225-1226	ILSALEIPDVGNL <u>S</u> QVVVMK	
1227-1234	SK	65
1261-1269	NDQGITNK	24
1274-1276	LVA <u>S</u> NWYNR	
1277-1296	<u>S</u> SR	
	TLGCSWEFIPVDDGWGERPL	105